Nonparametric estimation of the purity of a quantum state in quantum homodyne tomography with noisy data.

Katia Meziani

Nonparametric estimation of the purity of a quantum state in Quantum Homodyne Tomography with noisy data

Katia Méziani

email: meziani@math.jussieu.fr

Laboratoire de Probabilités et Modèles Aléatoires,

Université Paris VII (Denis Diderot),

case courier 7012, 2 Place Jussieu 75251 Paris cedex 05, France.

April 23, 2007

Abstract

The aim of this paper is to answer an important issue in quantum mechanics, namely to determine if a quantum state of a light beam is pure or mixed. The estimation of the purity is done from measurements by Quantum Homodyne Tomography performed on identically prepared quantum systems. The quantum state of the light is entirely characterized by the Wigner function, a density of generalized joint probability which can take negative values and which must respect certain constraints of positivity imposed by quantum physics. We propose to estimate a quadratic functional of the Wigner function by a kernel method as the physical measure of the purity of the state. We give also an adaptive estimator that does not depend on the smoothness parameters and we establish upper bound on the minimax risk over a class of infinitely differentiable functions.

AMS 2000 subject classifications: 62G05, 62G20, 81V80,
Key Words: Adaptive estimation, deconvolution, infinitely differentiable functions, minimax risk, quadratic functional estimation, quantum state, Wigner function, Radon transform, quantum homodyne tomography.

1 Introduction

In quantum mechanics, the quantum state of a system completely describes all aspects of the system. The instantaneous state of a quantum system encodes the probabilities of its measurable properties, or "observables" (examples of observables include energy, position, momentum and angular momentum). Generally, quantum mechanics does not assign deterministic values to observables. Instead, it makes predictions about probability distributions; that is, the probability of obtaining each of the possible outcomes from measuring an observable. In many applications of quantum information, one of the important elements which affect the result of quantum process, is the purity of quantum states produced or utilized. Hence, an interesting and important problem in quantum information is to estimate the purity of a quantum system [4, 30]. This problem is also strongly related to the estimation of the entanglement of multiparty systems [15, 1]. A state is called pure if it cannot be represented as a mixture (convex combination) of other states, i.e., if it is an extreme point of the convex set of states. All other states are called mixed states. The measurement technique of a quantum state is called Quantum Homodyne Tomography (QHT) and has been put in practice for the first time in [25]. We will detail this technique in Section 2.2. A quantum state is represented through two mathematical objects: the density matrix ρ and the associated real function of two variables W_ρ called the Wigner function [29].

In this paper we address the problem of estimating the quadratic functional $d^2 = \int W_\rho^2$ of the Wigner function of a monochromatic light in a cavity prepared in the state ρ by using QHT data measurement performed on independent, identical systems. Our model takes into account the detection losses occuring in the measurement, leading to an additional additive Gaussian noise. Our data consists of bivariate, independent, identically distributed observations in a double inverse Radon Transform (tomography) and convolution Gaussian random variable model that we describe in Section 2.4. The quantity $d^2 = \int W_\rho^2$ has an interest in itself as a physical measure of the purity of quantum state. It allows us to distinguish between pure state and mixed state since it always equals $\frac{1}{2\pi}$ in case of pure states (see Section 2.1 for relation between this quantity and the notion of purity) and is different from $\frac{1}{2\pi}$ if the state is mixed.
In general, W_ρ is regarded as a generalized joint probability density of the electric and magnetic fields of a laser beam, integrating to plus one over the whole plane. It does not satisfy all the properties of a proper probability density as it can, and normally does, go negative for states which have no classical model. It satisfies also certain intrinsic positivity constraints in the sense that it corresponds to a density matrix.

The problem of estimating quadratic functionals was studied in details in [6], where the problem of estimating the integral of the squared derivative of a probability density function was considered and nonparametric rates were obtained. These results have been extended in the density model on the estimation of general functionals of a density f of the type $\int f^2$ in [7], of the type $\int f^3$ in [17] and of the type $\int T(f)$ in [21] where minimax rate have been established. Minimax rates have also been obtained in [23] for the nonparametric estimation of $\|f\|_r$ in the classical white noise model. More recently, the estimation of $\int f^2$ in the convolution model have been treated in [9] for application to the goodness-of-fit test in L_2 distance.
The problem of adaptive estimation of general functionals in the white noise model

has been considered in [13] for $\int_0^1 f^2$, in [27] for $\int T(f)$ for arbitrary 4 times continuously differentiable functionals T and more rencently in [18] for sharp adaptive estimation of quadratic functionals.
In a positron emission tomography (PET) perspective, the problem of estimating a probability density from tomographic data at sharp minimax rates has been treated in [16] for bivariate density and in [12] for multi-dimensional density. Some functional estimation problems in the image model, like estimating the area of an image, have been considered in [19].
Quantum statistic models are more recent, the estimation of the Wigner function W_ρ has been treated in [14] in the case of ideal detection, that is without noise. The estimation of the Wigner function in our noisy model has been studied in a nonparametric framework in [10, 3] where minimax rate was established for the pointwise and the $\mathbb{L}_2$risk respectively.
We emphasize that in our paper we do not restrict ourselves to the parametric setting, but suppose that the Wigner function belongs to a nonparametric class $\mathcal{A}(\alpha, r, L, L')$ described in Section 2.4. We refer the interested reader to [2, 5] for further details on physical background.

In this paper we propose a kernel estimator for the quantity $d^2 = \int W_\rho^2$. We investigate the rate of convergence of the procedure and show that the bandwidth leading to the bias-variance trade-off depends on the parameter describing the functional class containing W_ρ. Therefore we propose an adaptive estimator based on a data-driven choice of the bandwidth. This adaptive estimator is shown to have the same rate of convergence. Let us briefly describe a possible application of our results to goodness-of-fit test in $\mathbb{L}_2$-norm in quantum statistics. The physical interpretation of such a test is to check whether the produced light pulse is in the known quantum state ρ_0 or not. This can be done via the Wigner functions as follows:

$$\begin{cases} H_0: & W_\rho = W_{\rho_0}, \\ H_1: & \sup_{W_\rho \in \mathcal{A}(\alpha,r,L,L')} \|W_\rho - W_{\rho_0}\|_2 \geq c \cdot \varphi_n \end{cases}$$

where φ_n is a sequence which tends to 0 when $n \to \infty$ and it is the testing rate and $\mathcal{A}(\alpha, r, L, L')$ is a class of smooth Wigner functions (see Section 2.4). We can device a test statistic based on the estimator of $d^2 = \int W_\rho^2$ constructed in this paper. Similary to [9] we conjecture that the testing rates are of the same order as the ones found in this paper.

The rest of the paper is organized as follows. In Sections 2.1 and 2.2, we make a short introduction to quantum mechanics. We formulate the statistical model in Section 2.4. In Section 3 we construct an estimator of the quadratic functional of the unknown Wigner function, and state a result on upper bound on the bias and the variance terms (proof in Section 4). Then we propose a choice of bandwidth independent of the smooth parameters yielding the same rate of convergence. Our main theoretical results are presented in Section 3.3. We present some examples of quantum states in Section 2.3.

2 Physical and statistical context

2.1 A short introduction to Quantum Mechanics

Quantum mechanics is a fundamental branch of theoretical physics, in the sense that it provides accurate and precise descriptions for many phenomena on the atomic and subatomic level. In the formalism of quantum mechanics, the state of a system at a given time is described by a complex wave function (sometimes referred to as orbitals in the case of atomic electrons), and more generally, elements of a complex vector space. Generally, quantum mechanics only makes predictions about probability distributions; that is, the probability of obtaining each of the possible outcomes from measuring an observable. Naturally, these probabilities will depend on the quantum state at the instant of the measurement. There are numerous mathematically equivalent formulations of quantum mechanics. Mathematically, the possible states of a quantum system are represented by unit vectors (called "state vectors") residing in the associated complex separable Hilbert space $\mathcal{H}$. In other words, the possible states are points in the projectivization of a Hilbert space. Each state is represented by a density matrix ρ which is a linear operator on the space $\mathcal{H}$ having the following properties:

- Self-adjoint (or Hermitian): $\rho = \rho^*$, where ρ^* is the adjoint of ρ.
- Positive: $\rho \geq 0$, or equivalently $\langle \psi, \rho\psi \rangle \geq 0$ for all $\psi \in \mathcal{H}$.
- Trace one: $Tr(\rho) = 1$.

A state is called pure if it cannot be represented as a mixture (convex combination) of other states, i.e., if it is an extreme point of the convex set of states. Thus, pure states are represented by one dimensional orthogonal projection operators i.e. $\rho = P_{\psi_j}$. All other states are called mixed states and for a separable Hilbert space $\mathcal{H}$, they can be expressed as

$$\rho = \sum_{i}^{dim\mathcal{H}} \rho_i P_{\psi_i}.$$

Due to the previously stated properties of ρ, $\rho_i \geq 0$ are the eigenvalues of ρ such that $\sum_i \rho_i = 1$, and P_{ψ_i} the projection onto the one dimensional space generated by the eigenvector $\psi_i \in \mathcal{H}$ of ρ.

Equivalently a corresponding Wigner function $W_\rho : \mathbb{R}^2 \to \mathbb{R}$ may be defined and describes completely the quantum state ρ. In general, W_ρ is regarded as a generalized joint probability density of two variables P and Q (the electric and magnetic fields of a laser beam). The Wigner function may take negative values but it integrates to plus one over the whole plane. It satisfies also certain intrinsic positivity constraints in the sense that it corresponds to a density matrix. (For further information on the Wigner function, we invite readers to consult the paper in [2].)

The important relation between ρ and W_ρ is the following one

$$2\pi \int_{\mathbb{R}^2} W_\rho^2(z)dz = Tr(\rho^2) = Tr(\sum_i \rho_i^2 P_{\psi_i}) = \sum_i \rho_i^2.$$

Then if the last sum $\sum_i \rho_i^2 = 1$, it means $\rho_i = \delta_{ij}$, thus $\rho = P_{\psi_j}$ is a pure state. We propose in this paper to study the quantity $\int_{\mathbb{R}^2} W_\rho^2(z)dz$ as a physical measure of purity.

2.2 Quantum Homodyne Tomography

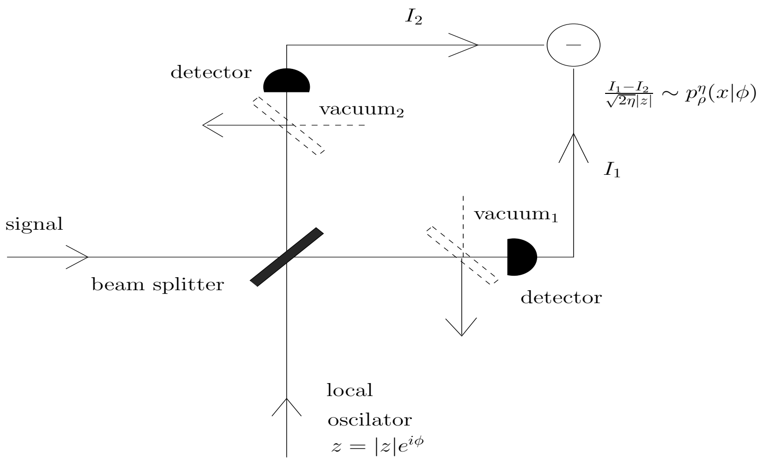

Figure 1: QHT mesurement

The theoretical foundation of quantum homodyne tomography was outlined in [28] and has inspired the first experiments determining the quantum state of a light field, initially with optical pulses in [25, 26]. The physicists developed a monochromatic laser in state ρ in a cavity. In order to study it, one takes measurement by QHT. This technique schematized in Figure 1 consists in mixing the light pulse in which we are interested with a laser of reference of high intensity $|z| >> 1$ called local oscillator. Just before the mixing the experimenter chooses the phase Φ of the local oscillator, randomly, uniformly distributed. After the mixing there are two emerging beams and each one is measured to give integrated currents I_1, I_2 proportional to the intensities. the effective result of the measurement is $X = \frac{I_2 - I_1}{|z|}$ which together with the phase Φ gives (X, Φ). It is widely admitted in the physical litterature (see [22]) that an additive gaussian noise is mixed with ideal data X, giving for known efficiency η, data Y.

2.3 Examples

Table 1 shows five examples of quantum pure states and one example of mixed state which can be created at this moment in laboratory. Among the pure states we consider the vacuum state which is the pure state of zero photons, the single photon state, the coherent state which characterizes the laser pulse with an average of N photons. The squeezed states (see e.g. [8]) have Gaussian Wigner functions whose variances in the two directions have a fixed product. The well-known Schrödinger Cat state is also a pure state discribed by a linear superposition of two coherent vectors (see e.g. [24]).

Table 1: Examples of quantum states

State	Fourier transform of Wigner function $\widetilde{W}_\rho(u,v)$	the Wigner function $W_\rho(p,q)$
Vacuum state	$\exp\left(\frac{-\|(u,v)\|_2^2}{4}\right)$	$\frac{1}{\pi}\exp(-q^2-p^2)$
Single photon state	$\left(1-\frac{\|(u,v)\|_2^2}{2}\right)\exp\left(\frac{-\|(u,v)\|_2^2}{4}\right)$	$\frac{1}{\pi}(2q^2+2p^2-1)\exp(-q^2-p^2)$
Schrödinger Cat $X_0>0$	$e^{\frac{-\|(u,v)\|_2^2}{4}}(\cos(2uX_0)$ $\left.+e^{-X_0^2}\cosh(X_0v)\right)/(2(1+e^{-X_0^2}))$	$e^{-p^2}\left(e^{-(q-X_0)^2}+e^{-(q+X_0)^2}\right.$ $\left.+2\cos(2pX_0)e^{-q^2}\right)/(2\pi(1+e^{-X_0^2}))$
Coherent state $N\in\mathbb{R}_+$	$\exp\left(\frac{-\|(u,v)\|_2^2}{4}+i\sqrt{N}v\right)$	$\frac{1}{\pi}\exp(-(q-\sqrt{N})^2-p^2)$
Squeezed state $N\in\mathbb{R}_+$, $\xi\in\mathbb{R}$	$\exp\left(-\frac{u^2}{4}e^{2\xi}-\frac{v^2}{4}e^{-2\xi}+iv\alpha\right)$	$\frac{1}{\pi}\exp(-e^{2\xi}(q-\alpha)^2-e^{-2\xi}p^2)$
Thermal state $\beta>0$	$\exp\left(\frac{-\|(u,v)\|_2^2}{4(\tanh(\beta/2))^2}\right)$	$\frac{\tanh(\beta/2)}{\pi}\exp(-(q^2+p^2)\tanh(\beta/2))$

Note that for pure states $d^2=\frac{1}{2\pi}$ and for the thermal state which is a mixed state $d^2 = \frac{\tanh(\beta/2)}{2\pi}$.

2.4 Problem formulation

The monochromatic laser in state ρ in a cavity is described by density matrices on the Hilbert space of complex valued square integrable functions on the line $\mathbb{L}_2(\mathbb{R})$. Those functions are called Wigner functions. In the present paper we estimate the integral of the square of the Wigner function from data measurement performed on n identical quantum systems.

Our statistical problem can been formulated as follows:

consider $(X_1, \Phi_1) \dots (X_n, \Phi_n)$ independent identically distributed random variables with values in $\mathbb{R} \times [0, \pi]$. The probability density of (X, Φ) equals the **Radon transform** $\Re[W_\rho]$ of the Wigner function with respect to the measure λ/π, where λ is the Lebesgue measure on $\mathbb{R} \times [0, \pi]$. Thus

$$p_\rho(x|\phi) := \Re[W_\rho](x, \phi) = \int_{-\infty}^{\infty} W_\rho(x\cos\phi + t\sin\phi, x\sin\phi - t\cos\phi)dt \qquad (1)$$

is the density of X given $\Phi = \phi$. As we annouced in the introduction we do not observe the ideal data (X_ℓ, Φ_ℓ) $\ell = 1, \dots, n$ but a degraded noisy version $(Y_1, \Phi_1) \dots (Y_n, \Phi_n)$, where

$$Y_\ell := \sqrt{\eta} X_\ell + \sqrt{(1-\eta)/2}\xi_\ell. \qquad (2)$$

Here ξ_ℓ are standard Gaussian random variables independent of all (X_k, Φ_k) and $0 < \eta < 1$ is a known parameter. The parameter η is called the detection efficiency and $1-\eta$ represents the proportion of photons which are not detected due to various losses in the measurement process. We denote $p_\rho^\eta(x, \phi)$ the density of (Y_ℓ, Φ_ℓ). Thus, $p_\rho^\eta(\cdot, \phi)$ is the convolution of the density $\frac{1}{\sqrt{\eta}} p_\rho(\frac{\cdot}{\sqrt{\eta}}, \phi)$ of (X_ℓ, Φ_ℓ) with the density of a centered Gaussian density having variance $(1-\eta)/2$. Let us define the following functional class $\mathcal{F}(\alpha, r, L)$:

$$\mathcal{F}(\alpha, r, L) = \left\{ f : \mathbb{R}^2 \to \mathbb{R}, \quad \int_{\mathbb{R}^2} |f(u,v)|^2 e^{2\alpha\|(u,v)\|^r} du dv \leqslant (2\pi)^2 L \right\},$$

where $0 < r \leq 2$, $\alpha > 0$, $L > 0$ and $\|(u,v)\| = \sqrt{u^2+v^2}$ is the euclidian norm. All the typical states ρ prepared in laboratory have density matrix with diagonal decreasing very fastly: for some $C > 0$, $B > 0$ and $r \in]0,2]$

$$|\rho_{m,\ell}| \leq C \exp(-B\alpha(m^{r/2} + \ell^{r/2})) \quad \forall m, \ell \in \mathbb{N}. \tag{3}$$

Recently, one has shown in [3] that quantum states satisfying (3) have Wigner function W_ρ in the class

$$\mathcal{A}(\alpha, r, L, L') = \left\{ W_\rho : \mathbb{R}^2 \to \mathbb{R}, \text{ Wigner function}, \ W_\rho \in \mathcal{F}(2^r\alpha, r, L'), \ \widetilde{W_\rho} \in \mathcal{F}(\alpha, r, L) \right\},$$

for some α, L, $L' > 0$ where $\widetilde{W_\rho}(u,v)$ denotes the Fourier transform of W_ρ w.r.t two variables. In this paper, we assume that the unknown Wigner function W_ρ belongs to the class $\mathcal{A}(\alpha, r, L, L')$ of infinitely differentiable functions.

3 Estimation procedure and main results

We are now able to define the estimation procedure of the quadratic functional $d^2 = \int W_\rho^2$ of the unknown function W_ρ based on data (Y_ℓ, ϕ_ℓ). Next we state an upper bound of the maximal risk uniformly over all Wigner functions in the class $\mathcal{A}(\alpha, r, L, L')$.

3.1 Kernel estimator

Let us define our estimator as a U-statistic of order 2:

Definition 1. *For any $\delta = \delta_n > 0$, we define the estimator*

$$d_n^2 = \frac{1}{n(n-1)} \sum_{j \neq k} \int_{\|z\| \leq 1/\delta} K_{\delta,n}\left([z, \Phi_j] - \frac{Y_j}{\sqrt{\eta}}\right) K_{\delta,n}\left([z, \Phi_k] - \frac{Y_k}{\sqrt{\eta}}\right) dz, \tag{4}$$

where

$$K_{\delta,n}(x) = \frac{1}{(2\pi)^2} \int_{-1/\delta}^{1/\delta} |t| e^{-itx} e^{\frac{t^2}{4}\frac{1-\eta}{\eta}} dt = \frac{1}{(2\pi)^2} \int_{-1/\delta}^{1/\delta} |t| \cos(tx) e^{\frac{t^2}{4}\frac{1-\eta}{\eta}} dt. \tag{5}$$

Note that the Fourier transform of $K_{\delta,n}$ is $\widetilde{K}_{\delta,n}(t) = \frac{1}{2\pi}|t|e^{\frac{t^2}{4}\frac{1-\eta}{\eta}}\mathbb{I}(|t| \leq 1/\delta)$, where $\mathbb{I}$ stands for the indicator function.

Let d_n^2 be the estimator defined by (4), having bandwidth $\delta > 0$. We call the bias and the variance of the estimator, respectively:

$$B(d_n^2) := |E_\rho[d_n^2] - d^2| \quad \text{and} \quad \mathrm{Var}(d_n^2) := E_\rho\left[|d_n^2 - E_\rho[d_n^2]|^2\right]. \tag{6}$$

3.2 Bias-variance decomposition

The following proposition plays an important role in the proof of the upper bound of the risk as we split it into the bias term and the variance term.

Proposition 1. *Let $a = \frac{1-\eta}{2\eta}$ and d_n^2 be the estimator defined by (4) with $\delta \to 0$ such that $e^{a/\delta^2}/(n\delta^2) \to 0$ as $n \to \infty$, then for all $\alpha > 0$, $L, L' > 0$ and $0 < r \leq 2$*

$$\sup_{W_\rho \in \mathcal{A}(\alpha,r,L,L')} B^2(d_n^2) \;\leq\; L^2 e^{-4\alpha/\delta^r}(1 + o(1)), \tag{7}$$

$$\sup_{W_\rho \in \mathcal{A}(\alpha,r,L,L')} Var(d_n^2) \;\leq\; \frac{32\pi L}{na^2\delta^2} e^{\frac{a}{\delta^2}}(1 + o(1)). \tag{8}$$

The proof of this proposition is given in Section 4.

3.3 Main results

Let d_n^2 be an estimator of $d^2 = \int W_\rho^2$ defined by (4). We measure the accuracy of d_n^2 by the maximal risk over the class $\mathcal{A}(\alpha, r, L, L')$

$$\mathcal{R}(d_n^2; \mathcal{A}(\alpha, r, L, L')) = \sup_{W_\rho \in \mathcal{A}(\alpha,r,L,L')} E_\rho[|d_n^2 - d^2|^2].$$

Here E_ρ and P_ρ denote the expected value and the probability when the true underlying quantum state is ρ.

Theorem 1. *Let* $(Y_\ell, \phi_\ell), \ell = 1, \dots, n$ *be i.i.d data coming from the model* (2) *where the underlying parameter is the Wigner function* W_ρ *lying in the class* $\mathcal{A}(\alpha, r, L, L')$, *with* $0 < r < 2$, $\alpha > 0$, $L, L' > 0$. *Let* $a = \frac{1-\eta}{2\eta}$, *then* d_n^2 *defined in* (4) *with bandwidth* $\delta := \delta_{opt}$ *chosen as the solution of the equation*

$$\frac{a}{\delta_{opt}^2} + \frac{4\alpha}{\delta_{opt}^r} = \log n - (\log\log n)^2, \tag{9}$$

satisfies the following upper bound

$$\limsup_{n\to\infty} \varphi_n^{-2} \mathcal{R}(d_n^2; \mathcal{A}(\alpha, r, L, L')) \le L^2, \tag{10}$$

where the rate of convergence is $\varphi_n^2 = e^{-4\alpha/\delta_{opt}^r}$.

Remark 1. *The previous theorem gives the upper bound of the risk. It is shown that the rate of convergence is given by the dominating term (bias term) at the selected bandwidth* $\delta := \delta_{opt}$. *Following the proof of the lower bound in [10], we can prove that similar lower bound holds in our setting when the Wigner function* W_ρ *belongs to the class* $\left\{W_\rho : \widetilde{W_\rho} \in \mathcal{F}(\alpha, r, L)\right\}$ *which is strictly larger than* $\mathcal{A}(\alpha, r, L)$. *Unfortunetly, the Wigner functions constructed in [10] for proving the lower bound do not belong to class* $\mathcal{F}(2^r\alpha, r, L')$.

Sketch of proof of Theorem 1 On the one hand, for $0 < r < 2$ and by (7) and (8), we select the bandwidth δ^* as

$$\delta^* = \arg\min_{\delta > 0} \left\{ \frac{C_V}{n\delta^2} e^{\frac{a}{\delta^2}} + C_B e^{-4\alpha/\delta^r} \right\},$$

by taking derivatives, δ^* is a positive real number satisfying

$$\frac{a}{\delta^{*2}} + \frac{4\alpha}{\delta^{*r}} = \log(\delta^{*4-r}) + \log n + const.$$

We notice that $B(d_n^2) \sim (\delta^*)^{r-2} Var(d_n^2)$, so the rate of convergence for the upper bound is given by the bias term of the estimator d_n^2 with $\delta = \delta^*$ i.e. $\varphi_n^2 = B(d_n^2)(1+$

$o(1)$). On the other hand, by taking $\delta := \delta_{opt}$ the unique solution of the equation

$$\frac{a}{\delta_{opt}^2} + \frac{4\alpha}{\delta_{opt}^r} = \log n - (\log\log n)^2,$$

the variance of the estimator d_n^2 with $\delta = \delta_{opt}$ is still smaller than its bias and its bias is of the same order as the bias of d_n^2 with optimal $\delta = \delta^*$ (see Lemma 8 in [11]). So, when replacing δ^* with the slightly modified δ_{opt} the upper bound of the minimax risk will remain asymptotically the same.

Remark If we consider the case $r \in]0,1]$, we can give a more explicit form for the bandwidth verifing (9) and thus, for the rate of convergence which is asymptotically equivalent to the bias term. Based on the results in [20], we make successive approximations starting with

$$\delta_0 := \left(\frac{\log n - (\log\log n)^2}{a}\right)^{-1/2},$$

and for all $k \geq 1$, if $r \in I_k =]\frac{2(k-1)}{k}, \frac{2k}{k+1}]$, we get recursively δ_k by plugging δ_{k-1} into $\delta_k = (\delta_0^{-2} - \frac{4\alpha}{a}\delta_{k-1}^{-r})^{-1/2}$. Then by choosing $\delta_{opt} = \delta_k$ and if $r \in I_k$, we obtain the following asymptotic equivalent of the rate of convergence

$$L^2 \exp\left(-4\alpha\delta_0^{-r} + C_1\delta_0^{2-r} - \ldots + C_{k-1}\delta_0^{2(k-1)-kr}\right).$$

Theorem 2. *Let $(Y_\ell, \phi_\ell), \ell = 1, \ldots, n$ be i.i.d data coming from the model (2) where the underlying parameter is the Wigner function W_ρ lying in the class $\mathcal{A}(\alpha, r, L, L')$, with $r = 2$, $\alpha > 0$, $L, L' > 0$. Let $a = \frac{1-\eta}{2\eta}$, then d_n^2 defined in (4) with bandwidth $\delta = \delta^* = \left(\frac{\log n - \log(\log n/(4\alpha+a))}{4\alpha+a}\right)^{-1/2}$ satisfies the following upper bound*

$$\limsup_{n\to\infty} \varphi_n^{-2}\mathcal{R}(d_n^2; \mathcal{A}(\alpha, r, L, L')) \leq C, \tag{11}$$

where the rate of convergence is $\varphi_n^2 = \left(\frac{n}{\log n}\right)^{-\frac{4\alpha}{4\alpha+a}}$, for some constant $C > 0$.

Sketch of proof of Theorem 2 For $r = 2$ and by (7) and (8), we select the bandwidth δ^* as

$$\widetilde{\delta} = \arg\min_{\delta>0} \left\{ \frac{C_V}{n\delta^2} e^{\frac{a}{\delta^2}} + C_B e^{-4\alpha/\delta^2} \right\},$$

by taking derivatives, we notice that $B(d_n^2) \sim Var(d_n^2)$ for $\delta = \widetilde{\delta}$ and that the rate of convergence is $\left(\frac{n}{\log n}\right)^{-\frac{4\alpha}{4\alpha+a}}$. It is easy to check that if we choose δ^* as bandwidth we get the same rate.

In the previous theorems, the bandwidth δ_{opt} depends on the parameters α and r of the class $\mathcal{A}(\alpha, r, L, L')$ which may be difficult to evaluate in practice. However, it is possible to construct an adaptive estimator which does not depend on these parameters and which has the same asymptotic behavior as in Theorem 1, provided that these parameters lie in a certain set. Note that the parameter η is supposed to be known. Define the set of parameters

$$\Theta_1 \quad = \quad \{(\alpha, r, L, L') : \alpha > 0, L, L' > 0, 0 < r < 1\}.$$

Theorem 3. *Let $(Y_\ell, \phi_\ell), \ell = 1, \dots, n$ be i.i.d data coming from the model (2). Let $d^2_{ad,n}$ be the estimator defined by*

$$d^2_{ad,n} = \frac{1}{n(n-1)} \sum_{j\neq k} \int_{\|z\|\leq 1/\delta_{ad}} K_{\delta_{ad},n}([z, \Phi_j] - \frac{Y_j}{\sqrt{\eta}}) K_{\delta_{ad},n}([z, \Phi_k] - \frac{Y_k}{\sqrt{\eta}}) dz,$$

with $\delta = \delta_{ad} = (\frac{\log n}{a} - 2\sqrt{\frac{\log n}{a}})^{-1/2}$. Then, for all $(\alpha, r, L, L') \in \Theta_1$,

$$\limsup_{n\to\infty} \sup_{W_\rho \in \mathcal{A}(\alpha,r,L,L')} E[|d^2_{\delta,n} - d^2|^2] \varphi_n^{-2} \leq 1,$$

where φ_n is the rate defined in Theorem 1.

Sketch of proof of Theorem 3 Over the set Θ_1, we easily check that, for $(\alpha, r, L, L') \in \Theta_1$

$$L^2 \exp\left(-\frac{4\alpha}{(\delta_{ad})^r}\right) \leq L^2 \exp\left(-\frac{4\alpha}{(\delta_{opt})^r}\right)(1 + o(1)),$$

thus the upper bound of the bias of $d^2_{ad,n}$ is not larger than the upper bound of the bias of d^2_n with $\delta = \delta_{opt}$. As $0 < r/2 < 1/2$ it is easy to remark that for n large enough $-(\frac{\log n}{a} - 2\sqrt{\frac{\log n}{a}})^{r/2} > -\frac{a}{4\alpha}\sqrt{\frac{\log n}{a}}$ and $\frac{1}{\delta^2_{ad}} \leq \frac{\log n}{a}$ and thus $\exp\left(-a\sqrt{\frac{\log n}{a}}\right) \leq \exp\left(-\frac{4\alpha}{(\delta_{ad})^r}\right)$. Then the dominating term in the variance found in (8)

$$\begin{aligned}
\frac{1}{n\delta^2_{ad}}\exp\left(\frac{a}{\delta^2_{ad}}\right) &\leq \left(\frac{\log n}{a}\right)\exp\left(-2a\sqrt{\frac{\log n}{2a}}\right) \\
&\leq \left(\frac{\log n}{a}\right)\exp\left(-a\sqrt{\frac{\log n}{2a}}\right)\exp\left(-a\sqrt{\frac{\log n}{2a}}\right) \\
&\leq \left(\frac{\log n}{a}\right)\exp\left(-a\sqrt{\frac{\log n}{2a}}\right)\exp\left(-\frac{4\alpha}{\delta^r_{ad}}\right) \\
&\leq o(1)\exp\left(-\frac{4\alpha}{(\delta_{opt})^r}\right).
\end{aligned}$$

Thus, $d^2_{ad,n}$ attains the rate φ^2_n.

4 Proof of Proposition 1

Most of the proofs make extensive use of the following equations and properties of Wigner functions. A remarkable relation links the Fourier transform of the Wigner function to the Fourier transform of its Radon transform. If we denote

$$\widetilde{W}_\rho(u,v) := \mathcal{F}_2[W_\rho](u,v),$$

then

$$\widetilde{W}_\rho(t\cos\phi, t\sin\phi) := \mathcal{F}_1[p_\rho(\cdot|\phi)](t) = E_\rho[e^{itX}|\Phi = \phi], \tag{12}$$

where $\mathcal{F}_2$, $\mathcal{F}_1$ denote the Fourier transform w.r.t two, respectively one variables. The Fourier transform w.r.t one variable of the density $p^\eta_\rho(\cdot|\phi)$ of Y when $\Phi = \phi$ is

$$\begin{aligned}
\mathcal{F}_1[p^\eta_\rho(\cdot|\phi)](t) &= \mathcal{F}_1[\frac{1}{\eta}p_\rho(\frac{\cdot}{\eta}|\phi)](t)\cdot\widetilde{N}^\eta(t) && (13)\\
&= \mathcal{F}_1[p_\rho(\cdot|\phi)](\sqrt{\eta}t)\cdot\widetilde{N}^\eta(t), && (14)
\end{aligned}$$

where $\widetilde{N}^\eta(t) = e^{-\frac{t^2}{4}(1-\eta)}$ denotes the characteristic function of the random variable $\sqrt{(1-\eta)/2}\xi \sim \mathcal{N}(0;(1-\eta)/2)$.

4.1 Proof of (7)

As (Y_k, Φ_k) and (Y_ℓ, Φ_ℓ) are i.i.d. for all $k \neq \ell$, we get

$$\begin{aligned} E[d_n^2] &= \frac{1}{n(n-1)} \sum_{j \neq k} \int_{\|z\| \leq 1/\delta} E[K_{\delta,n}([z,\Phi_j] - \frac{Y_j}{\sqrt{\eta}})] E[K_{\delta,n}([z,\Phi_k] - \frac{Y_k}{\sqrt{\eta}})] dz \\ &= \int_{\|z\| \leq 1/\delta} E[K_{\delta,n}([z,\Phi_1] - \frac{Y_1}{\sqrt{\eta}})] E[K_{\delta,n}([z,\Phi_2] - \frac{Y_2}{\sqrt{\eta}})] dz \\ &= \int_{\|z\| \leq 1/\delta} \left| E[K_{\delta,n}([z,\Phi] - \frac{Y}{\sqrt{\eta}})] \right|^2 dz. \end{aligned} \tag{15}$$

Moreover

$$\begin{aligned} E[K_{\delta,n}([z,\Phi] - \frac{Y}{\sqrt{\eta}})] &= \int_{\mathbb{R}} \int_0^\pi K_{\delta,n}([z,\phi] - \frac{y}{\sqrt{\eta}}) p_\rho^\eta(y,\phi) dy d\phi \\ &= \int_0^\pi \frac{1}{2\pi} \int_{\mathbb{R}} \mathcal{F}_1[K_{\delta,n} * \left(\sqrt{\eta} p_\rho^\eta(\cdot \sqrt{\eta}, \phi)\right)](t) e^{-it[z,\phi]} dt d\phi \\ &= \int_0^\pi \frac{1}{2\pi} \int_{\mathbb{R}} \widetilde{K}_{\delta,n}(t) \mathcal{F}_1[\sqrt{\eta} p_\rho^\eta(\cdot \sqrt{\eta}, \phi)](t) e^{-it[z,\phi]} dt d\phi. \end{aligned}$$

Then by using the expressions in (12), (13), (14) and a change of variables $(t\cos\phi, t\sin\phi) = w$, we get

$$\begin{aligned} & E[K_{\delta,n}([z,\Phi] - \frac{Y}{\sqrt{\eta}})] \\ &= \frac{1}{(2\pi)^2} \int_0^\pi \int_{\mathbb{R}} |t| e^{\frac{t^2}{4}\frac{1-\eta}{\eta}} \mathbb{I}(|t| \leq 1/\delta) \mathcal{F}_1[p_\rho(\cdot,\phi)](t) \widetilde{N}^\eta(t/\sqrt{\eta}) e^{-it[z,\phi]} dt d\phi \\ &= \frac{1}{(2\pi)^2} \int_0^\pi \int_{\mathbb{R}} |t| \, \mathbb{I}(|t| \leq 1/\delta) \mathcal{F}_1[p_\rho(\cdot,\phi)](t) e^{-it[z,\phi]} dt d\phi \\ &= \frac{1}{(2\pi)^2} \int_0^\pi \int_{\mathbb{R}} |t| \, \mathbb{I}(|t| \leq 1/\delta) \widetilde{W}_\rho(t\cos\phi, t\sin\phi) e^{-it[z,\phi]} dt d\phi \\ &= \frac{1}{(2\pi)^2} \int \mathbb{I}(\|w\| \leq 1/\delta) \widetilde{W}_\rho(w) e^{-i\langle z,w\rangle} dw. \end{aligned} \tag{16}$$

Thus, the Fourier transform of $E[K_{\delta,n}([z,\Phi]-\frac{Y}{\sqrt{\eta}})]$ as a function of z is $\widetilde{W}_\rho \cdot \mathbb{I}(\|\cdot\| \leq 1/\delta)$. We write then

$$E[K_{\delta,n}([z,\Phi]-\frac{Y}{\sqrt{\eta}})] \;=\; [W_\rho * K_\delta]\,(z), \tag{17}$$

where $\widetilde{K}_\delta(w) = \mathbb{I}(\|w\| \leq 1/\delta)$. Let us study the bias term by (6). By (15) and (17)

$$\begin{aligned}
&B(d_n^2)\\
&= \left|d^2 - E[d_n^2]\right| = \left|\int_{\mathbb{R}^2} W_\rho^2(z)dz - \int_{\|z\|\leq 1/\delta} \left|E[K_{\delta,n}([z,\Phi]-\frac{Y}{\sqrt{\eta}})]\right|^2 dz\right|\\
&= \left|\int_{\mathbb{R}^2} W_\rho^2(z)dz - \int_{\|z\|\leq 1/\delta}[W_\rho * K_\delta]^2(z)dz\right|\\
&\leq \left|\int_{\mathbb{R}^2} \left(W_\rho^2(z)dz - [W_\rho * K_\delta]^2(z)\right)dz\right| + \left|\int_{\|z\|>1/\delta}[W_\rho * K_\delta]^2(z)dz\right|\\
&\leq \left|\int_{\mathbb{R}^2} \left(W_\rho^2(z) - [W_\rho * K_\delta]^2(z)\right)dz\right| + 2\left|\int_{\|z\|>1/\delta} |[W_\rho * K_\delta](z) - W_\rho(z)|^2\,dz\right|\\
&\quad +2\left|\int_{\|z\|>1/\delta} |[W_\rho(z)|^2\,dz\right|\\
&= A_1 + 2A_2 + 2A_3,
\end{aligned} \tag{18}$$

where A_1, $2A_2$ and $2A_3$ are respectively the first, the second and the third term of the previous sum (18). By the Plancherel formula and using the smoothness of W_ρ

$$\begin{aligned}
A_1 &= \frac{1}{(2\pi)^2}\left|\int_{\mathbb{R}^2}\left(|\widetilde{W}_\rho(w)|^2 - |\widetilde{W}_\rho(w)\mathbb{I}(\|w\| \leq 1/\delta)|^2\right)dw\right|\\
&= \frac{1}{(2\pi)^2}\int_{\|w\|>1/\delta} |\widetilde{W}_\rho(w)|^2 dw \leq Le^{-\frac{2\alpha}{\delta^r}},
\end{aligned} \tag{19}$$

and, obviously,

$$\begin{aligned}
A_2 &= o(1)\int_{\mathbb{R}^2} |[W_\rho * K_\delta](z) - W_\rho(z)|^2\,dz\\
&= o(1)\frac{1}{(2\pi)^2}\int_{\mathbb{R}^2}\left|\widetilde{W}_\rho(w)\mathbb{I}(\|w\| \leq 1/\delta) - \widetilde{W}_\rho(w)\right|^2 dw\\
&= o(1)\frac{1}{(2\pi)^2}\int_{\|w\|>1/\delta} |\widetilde{W}_\rho(w)|^2 dw = o(1)Le^{-\frac{2\alpha}{\delta^r}},
\end{aligned} \tag{20}$$

as $\delta \to 0$, $n \to \infty$. Using now the asymptotic behaviour of W_ρ

$$A_3 \;=\; \int_{\|z\|>1/\delta} |[W_\rho(z)|^2\,dz \le (2\pi)^2 L e^{-\frac{2\alpha 2^r}{\delta^r}} = o(1)Le^{-\frac{2\alpha}{\delta^r}}, \tag{21}$$

as $\delta \to 0$, $n \to \infty$ and $2^r > 1$ $\forall r > 0$. Then, by using (19), (20) and (21)

$$B(d_n^2) \;\le\; Le^{-2\alpha/\delta^r}(1+o(1)), \,\text{as}\, \delta \to \infty.$$

4.2 Proof of (8)

We recall that $E[K_{\delta,n}([z,\Phi] - \frac{Y}{\sqrt{\eta}})] = [W_\rho * K_\delta](z)$, then

$$\begin{aligned}
& n(n-1)(d_n^2 - E[d_n^2]) \\
&= \sum_{j\neq k} \int_{\|z\|\le 1/\delta} \left\{ K_{\delta,n}([z,\Phi_j] - \frac{Y_j}{\sqrt{\eta}}) K_{\delta,n}([z,\Phi_k] - \frac{Y_k}{\sqrt{\eta}}) - [W_\rho * K_\delta]^2(z) \right\} dz \\
&= \sum_{j\neq k} \int_{\|z\|\le 1/\delta} \left[\left(K_{\delta,n}([z,\Phi_j] - \frac{Y_j}{\sqrt{\eta}}) - [W_\rho * K_\delta](z) \right) \right. \\
&\quad \times \left(K_{\delta,n}([z,\Phi_k] - \frac{Y_k}{\sqrt{\eta}}) - [W_\rho * K_\delta](z) \right) \\
&\quad + [W_\rho * K_\delta](z) \left\{ \left(K_{\delta,n}([z,\Phi_j] - \frac{Y_j}{\sqrt{\eta}}) - [W_\rho * K_\delta](z) \right) \right. \\
&\quad \left.\left. + \left(K_{\delta,n}([z,\Phi_k] - \frac{Y_k}{\sqrt{\eta}}) - [W_\rho * K_\delta](z) \right) \right\} \right] dz.
\end{aligned}$$

Then, $d_n^2 - E[d_n^2] = J_1 + J_2$ where we denote by J_1, J_2 respectively

$$\begin{aligned}
J_1 \;&=\; \frac{1}{n(n-1)} \sum_{j\neq k} \int_{\|z\|\le 1/\delta} \left(K_{\delta,n}([z,\Phi_j] - \frac{Y_j}{\sqrt{\eta}}) - [W_\rho * K_\delta](z) \right) \\
&\quad \times \left(K_{\delta,n}([z,\Phi_k] - \frac{Y_k}{\sqrt{\eta}}) - [W_\rho * K_\delta](z) \right) dz,
\end{aligned}$$

and

$$J_2 \;=\; \frac{2}{n} \sum_{\ell} \int_{\|z\|\le 1/\delta} [W_\rho * K_\delta](z) \left(K_{\delta,n}([z,\Phi_\ell] - \frac{Y_\ell}{\sqrt{\eta}}) - [W_\rho * K_\delta](z) \right) dz.$$

Then

$$Var(d_n^2) = E[(d_n^2 - E[d_n^2])^2] = E[J_1^2] + E[J_2^2] + 2E[J_1 J_2]. \tag{22}$$

See that the third part of the previous sum:

$$\begin{aligned}
& n^2(n-1)E[J_1 J_2] \\
= \;& 2\sum_{k\neq j}\sum_{\ell} E\left[\left\{\int_{\|z\|\leq 1/\delta}\left(K_{\delta,n}([z,\Phi_k] - \frac{Y_k}{\sqrt{\eta}}]) - [W_\rho * K_\delta](z)\right)\right.\right. \\
& \times \left.\left(K_{\delta,n}([z,\Phi_j] - \frac{Y_j}{\sqrt{\eta}}) - [W_\rho * K_\delta](z)\right) dz\right\} \\
& \times \left.\left\{\int_{\|z\|\leq 1/\delta}[W_\rho * K_\delta](z)\left(K_{\delta,n}([z,\Phi_\ell] - \frac{Y_\ell}{\sqrt{\eta}}) - [W_\rho * K_\delta](z)\right) dz\right\}\right] \\
= \;& 0,
\end{aligned}$$

by noticing $E[K_{\delta,n}([z,\Phi_\ell] - \frac{Y_\ell}{\sqrt{\eta}}) - [W_\rho * K_\delta](z)] = 0$ for all $\ell = 1, ..., n$ and because we always have either $\ell \neq k$ and $K_{\delta,n}([z,\Phi_\ell] - \frac{Y_\ell}{\sqrt{\eta}})$, $K_{\delta,n}([z,\Phi_k] - \frac{Y_k}{\sqrt{\eta}})$ are independent or $\ell \neq j$ and $K_{\delta,n}([z,\Phi_j] - \frac{Y_j}{\sqrt{\eta}})$, $K_{\delta,n}([z,\Phi_\ell] - \frac{Y_\ell}{\sqrt{\eta}})$ are independent. Now study

$$\begin{aligned}
& (n(n-1))^2 E[J_1^2] \\
= \;& E\left[\left(\sum_{j\neq k}\int_{\|z\|\leq 1/\delta}\left(K_{\delta,n}([z,\Phi_j] - \frac{Y_j}{\sqrt{\eta}}) - [W_\rho * K_\delta](z)\right)\right.\right. \\
& \times \left.\left.\left(K_{\delta,n}([z,\Phi_k] - \frac{Y_k}{\sqrt{\eta}}) - [W_\rho * K_\delta](z)\right) dz\right)^2\right] \\
= \;& \sum_{j_1\neq k_1}\sum_{j_2\neq k_2} E\left[\left\{\int_{\|z\|\leq 1/\delta}\left(K_{\delta,n}([z,\Phi_{j_1}] - \frac{Y_{j_1}}{\sqrt{\eta}}) - [W_\rho * K_\delta](z)\right)\right.\right. \\
& \times \left.\left(K_{\delta,n}([z,\Phi_{k_1}] - \frac{Y_{k_1}}{\sqrt{\eta}}) - [W_\rho * K_\delta](z)\right) dz\right\} \\
& \times \left\{\int_{\|z\|\leq 1/\delta}\left(K_{\delta,n}([z,\Phi_{j_2}] - \frac{Y_{j_2}}{\sqrt{\eta}}) - [W_\rho * K_\delta](z)\right)\right. \\
& \times \left.\left.\left(K_{\delta,n}([z,\Phi_{k_2}] - \frac{Y_{k_2}}{\sqrt{\eta}}) - [W_\rho * K_\delta](z)\right) dz\right\}\right].
\end{aligned}$$

Note that, as soon as either j_1 is different from k_2 and j_2, or k_1 is different from k_2 and j_2 the expected value is 0. Thus,

$$
\begin{aligned}
&(n(n-1))^2 E[J_1^2] \\
&= \sum_{j\neq k} E\left[\left(\int_{\|z\|\le 1/\delta}\left(K_{\delta,n}([z,\Phi_j]-\frac{Y_j}{\sqrt{\eta}})-[W_\rho * K_\delta](z)\right)\right.\right. \\
&\quad\left.\left.\times\left(K_{\delta,n}([z,\Phi_k]-\frac{Y_k}{\sqrt{\eta}})-[W_\rho * K_\delta](z)\right)dz\right)^2\right].
\end{aligned}
$$

By using the Cauchy Schwarz inequality, as (Y_k,Φ_k) and (Y_j,Φ_j) are i.i.d. and by the definition (5) of $K_{\delta,n}$, we get

$$
\begin{aligned}
&E[J_1^2] \\
&\le \frac{2}{n^2}\left(\int_{\|z\|\le 1/\delta} E\left[\left|K_{\delta,n}([z,\Phi]-\frac{Y}{\sqrt{\eta}})-[W_\rho * K_\delta](z)\right|^2\right]dz\right)^2 \\
&\le \frac{2}{n^2}\left(\int_{\|z\|\le 1/\delta} E\left[\left|K_{\delta,n}([z,\Phi]-\frac{Y}{\sqrt{\eta}})\right|^2\right]dz\right)^2 \\
&\le \frac{2}{n^2}\left(\int_{\|z\|\le 1/\delta}\|K_{\delta,n}\|_\infty^2\,dz\right)^2 \le \frac{2}{n^2}\left(\int_{\|z\|\le 1/\delta}\left(\int\left|\widetilde{K}_{\delta,n}(t)dt\right|\right)^2dz\right)^2 \\
&= \frac{2}{n^2}\left(\int_{\|z\|\le 1/\delta}\left(\int_{|t|\le 1/\delta}|t|e^{t^2\frac{1-\eta}{4\eta}}dt\right)^2dz\right)^2 \\
&\le \frac{8\pi^2}{n^2\delta^4}\left(\frac{4\eta}{1-\eta}\right)^4 e^{\frac{1-\eta}{\eta\delta^2}}(1+o(1)). \qquad (23)
\end{aligned}
$$

The term $E[J_2^2]$ can be bounded as follows

$$
\begin{aligned}
&E[J_2^2] \\
&= E\left[\left(\frac{2}{n}\sum_\ell\int_{\|z\|\le 1/\delta}[W_\rho * K_\delta](z)\left(K_{\delta,n}([z,\Phi_\ell]-\frac{Y_\ell}{\sqrt{\eta}})-[W_\rho * K_\delta](z)\right)\right)^2\right] \\
&= \frac{4}{n^2}\sum_\ell E\left[\left(\int_{\|z\|\le 1/\delta}[W_\rho * K_\delta](z)\left(K_{\delta,n}([z,\Phi_\ell]-\frac{Y_\ell}{\sqrt{\eta}})-[W_\rho * K_\delta](z)\right)dz\right)^2\right] \\
&\le \frac{4}{n}E\left[\left(\int_{\|z\|\le 1/\delta}[W_\rho * K_\delta](z)K_{\delta,n}([z,\Phi]-\frac{Y}{\sqrt{\eta}})dz\right)^2\right].
\end{aligned}
$$

By Cauchy Schwarz inequality

$$
\begin{aligned}
&E[J_2^2] \\
&\leq \frac{4}{n}\int_{\|z\|\leq 1/\delta}\left|E[K_{\delta,n}([z,\Phi]-\frac{Y}{\sqrt{\eta}})]\right|^2 dz \int_{\|z\|\leq 1/\delta} E\left[\left|K_{\delta,n}([z,\Phi]-\frac{Y}{\sqrt{\eta}})\right|^2\right] dz \\
&\leq \frac{4}{n}\int_{\|z\|\leq 1/\delta} |[W_\rho * K_\delta](z)|^2\, dz \int_{\|z\|\leq 1/\delta}\left(\int_{|t|\leq 1/\delta} |t| e^{t^2\frac{1-\eta}{4\eta}}\, dt\right)^2 dz \\
&\leq \frac{8\pi}{n\delta^2}\left(\frac{4\eta}{1-\eta}\right)^2 e^{\frac{1-\eta}{2\eta\delta^2}}(1+o(1))\int_{\mathbb{R}^2} |[W_\rho * K_\delta](z)|^2\, dz.
\end{aligned}
$$

Parseval's equality and the asymptotic behaviour of $\widetilde{W}_\rho$ yield

$$
\begin{aligned}
&E[J_2^2] \\
&\leq \frac{8\pi}{n\delta^2}\left(\frac{4\eta}{1-\eta}\right)^2 e^{\frac{1-\eta}{2\eta\delta^2}}(1+o(1))\frac{1}{(2\pi)^2}\int_{\mathbb{R}^2} \mathbb{I}(\|w\|\leq 1/\delta)\left|\widetilde{W}_\rho(w)\right|^2 dw \\
&\leq \frac{8\pi}{n\delta^2}\left(\frac{4\eta}{1-\eta}\right)^2 e^{\frac{1-\eta}{2\eta\delta^2}}(1+o(1))\frac{1}{(2\pi)^2}\int_{\mathbb{R}^2} \left|\widetilde{W}_\rho(w)\right|^2 e^{2\alpha\|w\|^r} dw \\
&\leq \frac{8\pi L}{n\delta^2}\left(\frac{4\eta}{1-\eta}\right)^2 e^{\frac{1-\eta}{2\eta\delta^2}}(1+o(1)). \qquad (24)
\end{aligned}
$$

In view of (23), (24) we have $E[J_1^2] \sim E^2[J_2^2]$. As one chooses δ such that the variance term tends to 0, we conclude by using (22).

References

[1] L. Aolita. Measuring multipartite concurrence with a single factorizable observable. *Physical Review Letters*, 97(050501), 2006.

[2] L. M. Artiles, R. D. Gill, and M. I. Guţă. An invitation to quantum tomography. *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, 67(1):109–134, 2005.

[3] J.M. Aubry, C. Butucea, and M. Guţă. $\mathbb{L}_2$ estimation of the Wigner function in quantum homodyne tomography with noisy data. Manuscript.

[4] E. Bagan, M. Ballester, R. Gill, A. Monras, and R. Munoz-Tapia. Optimal full estimation of qubit mixed states. *Physical Review*, A 73(032301), 2006.

[5] O. E. Barndorff-Nielsen, R. D. Gill, and P. E. Jupp. On quantum statistical inference. *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, 65(4):775–816, 2003. With discussion and reply by the authors.

[6] P. J. Bickel and Y. Ritov. Estimating integrated squared density derivatives: sharp best order of convergence estimates. *Sankhyā (Statistics). The Indian Journal of Statistics. Series A*, 50(3):381–393, 1988.

[7] L. Birgé and P. Massart. Estimation of integral functionals of a density. *The Annals of Statistics*, 23(1):11–29, 1995.

[8] G. Breitenbach, S. Schiller, and J. Mlynek. Measurement of the quantum states of squeezed light. *Nature*, 387:471–475, 1997.

[9] C. Butucea. Goodness-of-fit testing and quadratic functional estimation from indirect observations. *The Annals of Statistics*, to appear, 2007.

[10] C. Butucea, M. Guţă, and L. M. Artiles. Minimax and adaptive estimation of the Wigner function in quantum homodyne tomography with noisy data. *The Annals of Statistics*, to appear, 2007.

[11] C. Butucea and A. Tsybakov. Sharp optimality for density deconvolution with dominating bias. *Theory of Probability and its Applications*, to appear, 2007.

[12] L. Cavalier. Efficient estimation of a density in a problem of tomography. *The Annals of Statistics*, 28(2):630–647, 2000.

[13] S. Efromovich and M. Low. On optimal adaptive estimation of a quadratic functional. *The Annals of Statistics*, 24(3):1106–1125, 1996.

[14] M. Guţă and L. Artiles. Minimax estimation of the Wigner in quantum homodyne tomography with ideal detectors. *Mathematical Methods of Statistics*, to appear, 2007.

[15] P. Horodecki. Measuring quantum entanglement without prior state reconstruction. *Physical Review Letters*, 90(167901), 2003.

[16] I. M. Johnstone and B. W. Silverman. Speed of estimation in positron emission tomography and related inverse problems. *The Annals of Statistics*, 18(1):251–280, 1990.

[17] G. Kerkyacharian and D. Picard. Estimating nonquadratic functionals of a density using Haar wavelets. *The Annals of Statistics*, 24(2):485–507, 1996.

[18] J. Klemelä. Sharp adaptive estimation of quadratic functionals. *Probability Theory Related Fields*, 134(4):539–564, 2006.

[19] A. P. Korostelëv and A. B. Tsybakov. *Minimax theory of image reconstruction*, volume 82 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1993.

[20] C. Lacour. Rates of convergence for nonparametric deconvolution. *Comptes Rendus Mathématique. Académie des Sciences. Paris*, 342(11):877–882, 2006.

[21] B. Laurent. Efficient estimation of integral functionals of a density. *The Annals of Statistics*, 24(2):659–681, 1996.

[22] U. Leonhardt. *Measuring the Quantum State of Light*. Cambridge University Press, 1997.

[23] O. Lepski, A. Nemirovski, and V. Spokoiny. On estimation of the L_r norm of a regression function. *Probability Theory Related Fields*, 113(2):221–253, 1999.

[24] A. Ourjoumtsev, R. Tualle-Brouri, J. Laurat, and P. Grangier. Generating optical schrödinger kittens for quantum information processing. *Science*, 312(5770):83–86, 2006.

[25] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani. Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum. *Physical Review Letters*, 70(9):1244–1247, March 1993.

[26] D.J. Smithey, M. Beck, M.J. Cooper, and M. G. Raymer. Experimental determination of number-phase uncertainty relations. *Optics Letters*, 18:1259–1261, 1993.

[27] K. Tribouley. Adaptive estimation of integrated functionals. *Mathematical Methods of Statistics*, 2000.

[28] K. Vogel and H. Risken. Determination of quasiprobabiltity distributions in terms of probability distributions for rotated quadrature phase. *Physical Review*, A(40):2847–2849, March 1989.

[29] E. Wigner. On the quantum correction for thermodynamic equations. *Physical Review*, 40(4):749–759, 1932.

[30] C. Zhang, G. Wang, and M. Ying. Discrimination between pure states and mixed states. preprint arXiv:quant-ph/0611122.